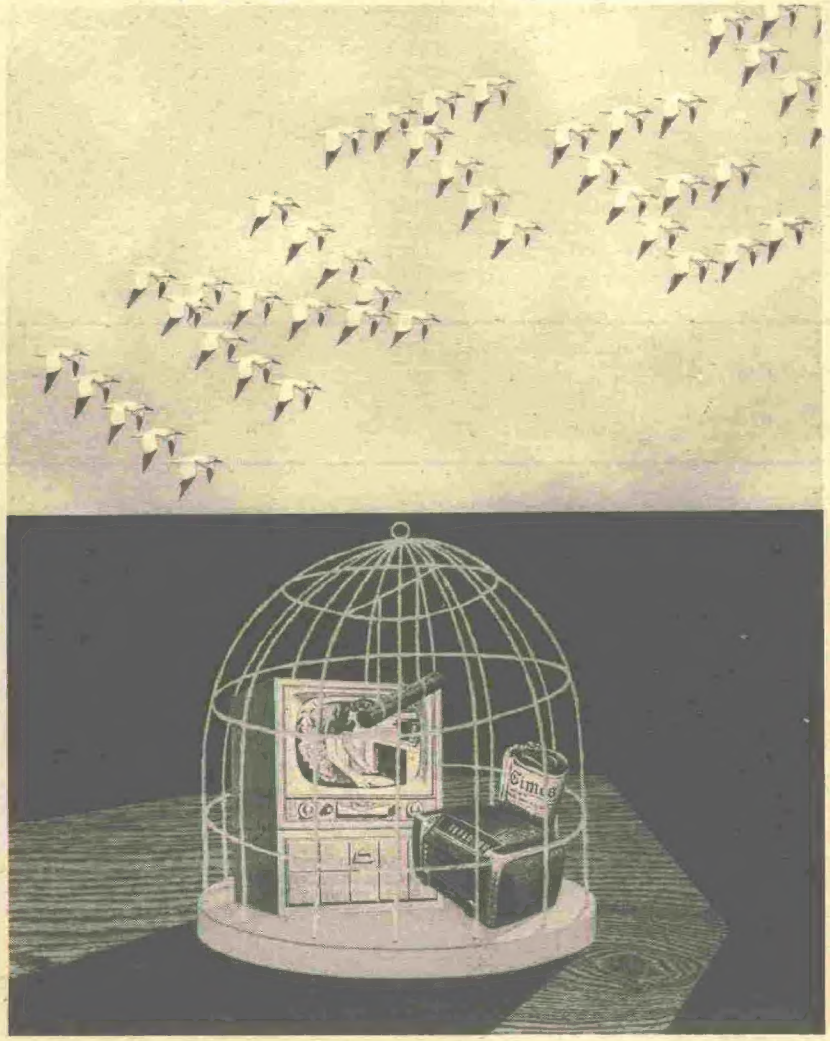
สุรศักดิ์ สงวนพงษ์

ตามล่าขุมข้อมูล อินเทอร์เน็ต

แยกประเภทของบริการค้นหาข้อมูลแบบออนไลน์ในอินเทอร์เน็ต เพื่อมุ่งหวังให้มีการใช้ทรัพยากรอย่างคุ้มค่า และเรียกค้นด้วยบริการที่เหมาะสม



ไม่ว่าครั้งแรกที่เริ่มทดลองใช้อินเทอร์เน็ตเพื่อค้นหาข้อมูลเรื่องใดเรื่องหนึ่ง มักต้องใช้เวลาอย่างมากเพื่อสืบเสาะหาแหล่งข้อมูล หรืออาจพบปัญหาตั้งแต่ว่าจะเลือกใช้บริการแบบใดสำหรับค้นหาข้อมูลที่ต้องการ เนื่องด้วยทรัพยากรข้อมูลในอินเทอร์เน็ตมีหลากหลายและกระจัดกระจายอย่างไม่เป็นหมวดหมู่ ในอินเทอร์เน็ตก็มีเครื่องมือหลากประเภท การเลือกใช้เครื่องมือ การใช้บริการอย่างถูกต้องและมีประสิทธิภาพ จึงต้องอาศัยประสบการณ์การทดลองใช้และความคุ้นเคยเป็นสำคัญ

เลือกใช้เครื่องมืออะไรดี

จากปรัชญาของการวางเครือข่ายซึ่งมุ่งหวังให้มีการใช้ทรัพยากรอย่างคุ้มค่า เครือข่ายอินเทอร์เน็ตจึงเปิดให้บริการข้อมูลโดยไม่คิดมูลค่า และสามารถเรียกค้นด้วยบริการที่เหมาะสม หากจะแยกประเภทของบริการค้นหาข้อมูลแบบออนไลน์ในอินเทอร์เน็ตแล้วสามารถแบ่งออกได้ดังต่อไปนี้

กลุ่มข่าวยูสเน็ต

ผู้ใช้เครือข่ายอินเทอร์เน็ตจำนวนมากทั่วโลกได้รวมตัวกันเป็นกลุ่มเพื่อการแลกเปลี่ยนความคิดเห็นในหัวข้อต่าง ๆ ผ่านทางเครือข่าย ซึ่งรู้จักกันในชื่อของ ยูสเน็ต (Usenet)

บริการยูสเน็ตเป็นจุดเริ่มต้นที่ดีสำหรับการค้นหาข้อมูลในเรื่องใดเรื่องหนึ่ง เนื่องจากในยูสเน็ตมีการจัดแบ่งหัวข้ออย่างเป็นระบบดังตารางที่ 1

ตารางที่ 1 กลุ่มข่าวหลักในยูสเน็ต

กลุ่ม	หัวข้อเรื่อง
comp	คอมพิวเตอร์ (computer)
misc	ข่าวที่ไม่สามารถจัดลงในกลุ่มใด (miscelleneous)
news	เกี่ยวกับยูสเน็ตเอง (news)
rec	สันทนาการ ศิลปะ (recreation)
sci	วิทยาศาสตร์ทุกแขนง (science)
soc	สังคม วัฒนธรรม (social)
talk	กลุ่มข่าวสำหรับการอภิปราย (talk)

หากต้องการค้นหาข้อมูลในเรื่องใดก็เพียงแต่เข้าไปอ่านสายการสนทนาในกลุ่มข่าวนั้น ๆ หรือค้นหาหัวข้อข่าวที่สมาชิกในกลุ่มได้จัดเตรียมเป็นสรุปคำถามคำตอบ (FAQ หรือ Frequently Ask Questions) ไว้ให้แล้ว ตัวอย่างในรูปที่ 1 แสดงหัวข้อ FAQ ของกลุ่มข่าวทางด้านไวรัสวิทยา (bionet.virology)

นอกจากการค้นหาในกลุ่มข่าวแต่ละกลุ่มโดยตรงแล้ว ยังมีกลุ่มข่าว news.answers ที่ใช้ส่ง FAQ จากหลาย ๆ กลุ่มข่าวมารวมไว้ที่นี่ FAQ ในกลุ่มนี้จะมีการปรับปรุงอยู่เป็นระยะ ดังรูปที่ 2

```
          bionet.virology (15T 51A OK 0H R)                    h=help
17   +     mAb to EBV glycoprotein B homolog                   David Koelle
18   +     monoclonal antibodies against reovirus ty           Toshiro hamamoto
19   + 4   Ebola/Hiv multiple questions                        Mari Noerr
20   + 5   Ebola Virus Question                                RGALLA41@MAINE.MAI
21   +     BIOSCI miniFAQ, ver. 14-DEC-95                      BIOSCI Administrat
22   + 2   Hi!                                                 Alexandre Perusse
23   + 3   Computer Virus Analogy                              Peter Leaback
24   +     International Herpesvirus Workshop                  Wolfram Brune
25   +     Points of interest                                  Thomas Nixon
26   +     postdoc. position available                         Linkman
27   +     NEED HELP WITH PENDIMETHALIN (PROWL(R)) HERBI       Kurt Sepp Deifel
<n>=set current to n, TAB=next unread, /=search pattern, ^K)ill/select,
a)uthor search, c)atchup, j=line down, k=line up, K=mark read, l)ist thread,
|=pipe, m)ail, o=print, q)uit, r=toggle all/unread, s)ave, t)ag, w=post
```

รูปที่ 1 ข่าวลำดับที่ 21 เป็น FAQ ประจำกลุ่มข่าว bionet.virology

```
          news.answers (1557T 1702A 0K 0H R)                   h=help
 737          Dreams FAQ Pt.3/4: About Lucid Dreaming          Olaf Titz
 738   +      Dreams FAQ Pt.4/4: Research, Help, reading re    Olaf Titz
 739   +      rec.music.makers.piano FAQ-General Topics        Isako Hoshino
 740   +      rec.music.makers.piano FAQ-Playing from Memor    Isako Hoshino
 741   +      rec.music.artists.debbie-gibson FAQ v3.0         Levien de Braal
 742   +      rec.music.makers.piano FAQ-Piano Purchase and    Isako Hoshin
 743   +      Judaism Reading List: Chasidism (Pt. VIII)       Daniel Faigin
 744   + 2    Cosmic Encounter Frequently Asked Questions      Kurt Adam
 745   +      alt.music.james-taylor FAQ                       Joel Risberg
 746   +      List of Medical FAQ Postings (1/2)               Bruce C McKenzie
 747   +      List of Medical FAQ Postings (2/2)               Bruce C McKenzie
 748   +      soc.genealogy.misc Frequently Asked Questions    john@meertech.demo
 749   +      Acorn ftp and mail-server archives (fortnight    Gerben Vos
 750   +      Misc.FS+Biz.Mktplc ADVERTISING FAQ-INFO FOR      dank@metrics.com
 751   +      Misc.FS+Biz.Mktplc TRANSACTIONS FAQ-INFO FOR     Dan King
 752   +      PC 3D Graphics Accelerators FAQ (Part 1 of 2)    Blair MacIntyre
<n>=set current to n, TAB=next unread, /=search pattern, ^K)ill/select,
 a)uthor search, c)atchup, j=line down, k=line up, K=mark read, l)istthread,
|=pipe, m)ail, o=print, q)uit, r=toggle all/unread, s)ave, t)ag, w=post
```

รูปที่ 2 ตัวอย่าง FAQ ในหัวข้อต่าง ๆ จาก news.answers

บริการ FTP

ข้อมูลที่สามารถถ่ายโอนด้วย FTP มักเป็นโปรแกรมแชร์แวร์หรือฟรีแวร์ในระบบปฏิบัติการยูนิกซ์ พีซี หรือแมคอินทอช ในขณะเดียวกันกลับมีข้อมูลทั่วไป เช่น ข่าวประจำวัน หรือบทความอีกเป็นจำนวนมาก

การใช้ FTP จึงต้องทราบชื่อเครื่องที่ให้บริการ หากไม่ทราบว่าศูนย์บริการใดมีข้อมูลประเภทใดบ้างย่อมเป็นเรื่องยาก เพราะการทดลองสุ่มค้นในศูนย์บริการแห่งใดแห่งหนึ่งจะเสียเวลาอย่างมาก หนทางที่ช่วยในการสืบค้นข้อมูลได้ง่ายขึ้นคือตรวจจากหัวข้อข่าว "Anonymous FTP: Sitelist" จากกลุ่มข่าว news.answers ดังรูปที่ 3 ซึ่งมีการเก็บรวบรวมรายชื่อศูนย์บริการและข้อมูลที่เก็บอยู่ หรือต่อเชื่อมไปค้นหารายชื่อศูนย์บริการด้วยอาร์ชี

โกเฟอร์

โกเฟอร์ (Gopher) เป็นโปรแกรมประยุกต์ที่ใช้สำหรับการเปิดค้นหาข้อมูลและขอใช้บริการด้วยระบบเมนู โปรแกรมโกเฟอร์เป็นเสมือนคลังห้องสมุดและเป็นจุดศูนย์รวมการเรียกใช้บริการต่าง ๆ ที่มีอยู่ในอินเทอร์เน็ตได้อย่างสะดวกและง่ายดาย

โกเฟอร์ทำงานด้วยระบบเมนูโดยให้ผู้ใช้เลือกค้นหาข้อมูลไปทีละหัวข้อ และอาจมีเมนูย่อยให้เลือกต่อไป เมื่อเลือกไปจนกระทั่งถึงเมนูชั้นในสุดของหัวข้อนั้น ๆ โกเฟอร์ก็จะแสดงข้อมูลบนจอภาพให้พลิกอ่านไปทีละหน้า

บริการในโกเฟอร์ไม่ได้แยกหมวดหมู่ข้อมูล ผู้ใช้จึงต้องอาศัยการทดลองสุ่มไปตามเมนูซึ่งลำบากต่อการค้นหา ด้วยเหตุนี้จึงมีการจัดทำดรรชนีช่วยค้นด้วยการป้อนคำสำคัญและเรียกระบบนี้ว่า Veronica (Very Easy Rodent-Oriented Netwide Index to Computerized Archives)

การใช้ Veronica จะเลือกจากเมนูที่มีหัวข้อว่า search by titles หรือคำอื่นที่ใกล้เคียงกัน รายการที่ 3 ในรูปที่ 4 เป็นหัวข้อการเข้าไปสู่ Veronica โดยใช้บริการโกเฟอร์จากมหาวิทยาลัยแห่งรัฐมิชิแกน

```
     news.answers (1557T 1702A 0K 0H R)                          h=help
 81  +    COMP.EMULATORS.CBM: Emulation FAQ for Commodo    'tsr'
 82  +    COMP.EMULATORS.CBM: Emulation FAQ for Commodo    'tsr'
 83  +    COMP.EMULATORS.CBM: Emulation FAQ for Commodo    'tsr'
 84  +    COMP.EMULATORS.CBM: Emulation FAQ for Commodo    'tsr'
 85  + 2  alt.fan.douglas-adams FAQ                        Nathan Hughes
 86  +    rec.pets.cats:  Ragdolls  Breed-FAQ              Jill Keating
 87  +    rec.pets.cats:  Traditional Siamese  Breed-FA    Laura Gilbreath
 88  +    Anonymous FTP: Frequently Asked Questions (FA    Perry Rovers
 89  +    Anonymous FTP: Sitelist Part 3 of 23     [03/2   Perry Rovers
 90  +    Anonymous FTP: Sitelist Part 4 of 23     [04/2   Perry Rovers
 91  +    Anonymous FTP: Sitelist Part 5 of 23     [05/2   Perry Rovers
 92  +    Anonymous FTP: Sitelist Part 9 of 23     [09/2   Perry Rovers
 93  +    Anonymous FTP: Sitelist Part 10 of 23    [10/2   Perry Rovers
 94  +    Anonymous FTP: Sitelist Part 19 of 23    [19/2   Perry Rovers
 95  +    Anonymous FTP: Sitelist Part 20 of 23    [20/2   Perry Rovers
 96  +    Usenet Info Center FAQ                           Kevin Atkinson
 <n>=set current to n, TAB=next unread, /=search pattern, ^K)ill/select,
 a)uthor search, c)atchup, j=line down, k=line up, K=mark read, l)ist thread,
|=pipe, m)ail, o=print, q)uit, r=toggle all/unread  s)ave, t)ag,w=post
```

รูปที่ 3
ข่าวลำดับที่ 88-95 เป็นส่วนหนึ่งของรายชื่อศูนย์ FTP ทั่วโลกและข้อมูลที่เก็บอยู่

```
                Internet Gopher Information Client v1.13

                    Root gopher server: gopher.msu.edu

      1.  Gopher at Michigan State University.
      2.  Help Using Gopher (More About Gopher)/
->    3.  Keyword Search of Titles in MSU's Gopher <?>
      4.  About Michigan State University/
      5.  MSU Campus Events/
      6.  News & Weather/
      7.  Phone Books & Other Directories/
      8.  Information for the MSU Community/
      9.  Computing & Technology Services/
      10. Libraries/
      11. MSU Services & Facilities/
      12. Outreach / Extension / Community Affairs/
      13. Network & Database Resources/
```

รูปที่ 4
เมนูโกเฟอร์ที่ให้บริการ Veronica

เวย์ส

หากเปรียบเทียบเวย์สกับเครื่องมืออื่น ๆ ในอินเทอร์เน็ตเช่นอาร์ชีซึ่งเป็นบริการค้นหาแหล่งที่เก็บแฟ้มข้อมูลโดยผู้ใช้ต้องป้อน ชื่อ แฟ้มที่ต้องการค้นหา ในขณะที่โกเฟอร์เป็นการค้นข้อมูลตามระบบเมนู แต่เวย์สเป็นบริการข้อมูลตามชื่อเรื่องหรือชื่อเนื้อหาที่ต้องการ

โปรแกรมเวย์สจะพยายามค้นเอกสารที่คิดว่าเกี่ยวข้องมากที่สุดกับคำหรือวลีสำคัญที่ผู้ใช้กำหนด แหล่งข้อมูลหรือฐานข้อมูลที่ต่อเชื่อมกันในระบบเวย์สประกอบด้วยแฟ้มข้อมูลจำนวนมหาศาล

ฐานข้อมูลที่ต่อเชื่อมอยู่กับระบบเวย์สในปัจจุบันนี้มีหลายร้อยแห่งที่ให้บริการข้อมูลโดยไม่คิดมูลค่า หากแต่กรรมวิธีในการค้นข้อมูลด้วยเวย์สจะไม่สะดวกเหมือนบริการอื่น

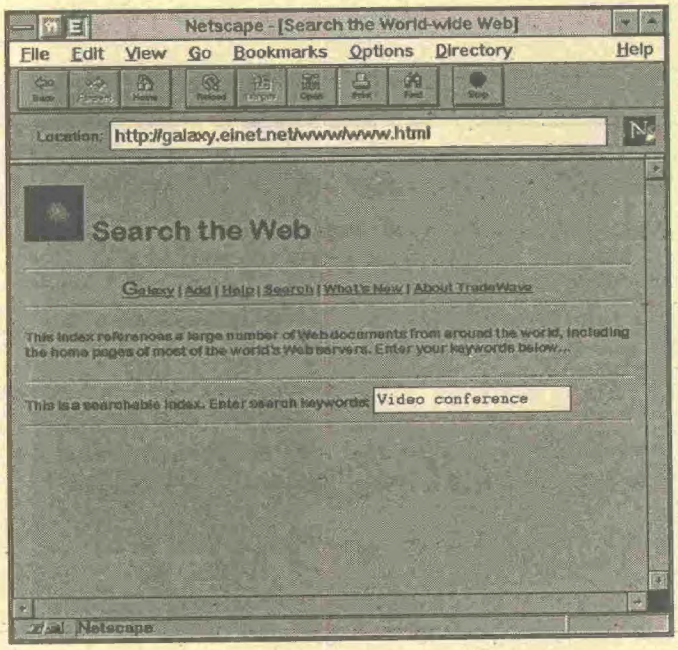
WWW

World-Wide-Web หรือ WWW นับเป็นบริการข้อมูลข่าวสารในอินเทอร์เน็ตที่ได้รับความนิยมสูงสุดในปัจจุบัน เนื่องจากเป็นเครื่องมือที่ง่ายต่อการใช้งาน และได้ผนวกบริการข้อมูลอื่นรวมไว้ในตัว เช่น การโอนย้ายแฟ้ม โกเฟอร์ และยูสเน็ต เป็นต้น

ในปัจจุบันมีศูนย์ WWW ทั่วโลกจำนวนมาก แต่ละศูนย์บริการจะมีข้อมูลเฉพาะเรื่อง จึงได้มีการจัดทำระบบสืบค้นข้อมูล (Search engines) ว่าอยู่ที่ศูนย์ใดบ้าง ดังตารางที่ 2

ตารางที่ 2 ศูนย์ WWW ที่มีดรรชนีหรือไดเรคทอรีให้บริการค้นหาข้อมูล

ศูนย์	URL
Aliweb	http://.web.nexor.co.uk/public/aliweb/aliweb.html
Excite	http://www.excite.com/
InfoSeek	http://www.infoseek.com/
Lycos	http://lycos.cs.cmu.edu/
New Riders' office WWW yellow pages	http://www.mcp.com/newriders/wwwyp/index.html
WebCrawler	http://webcrawler.com/
World Wide Web Worm	http://www.cs.colorado.edu/home/mcbryan/WWW.html
Yahoo	http://www.yahoo.com/



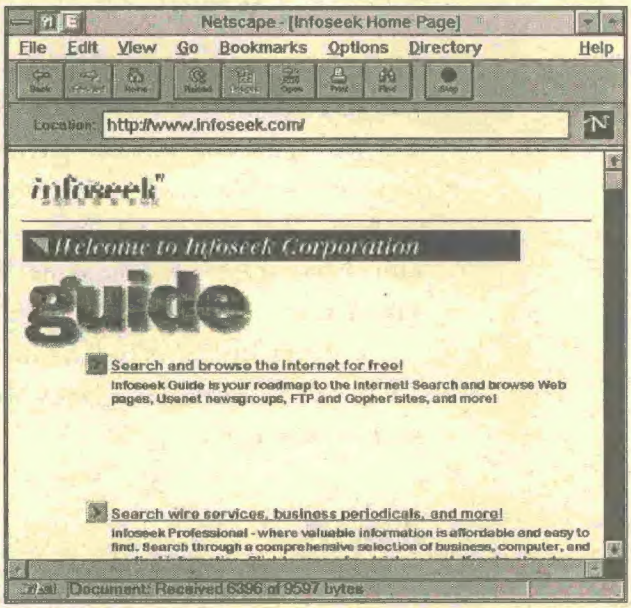
รูปที่ 5
ศูนย์ EINET Galaxy

- **EINET Galaxy** เป็นศูนย์บริการค้นหาที่จัดทำดรรชนีศูนย์ WWW ไว้ และเมื่อป้อนหัวข้อเรื่องที่ต้องการจะได้ผลลัพธ์ออกมาเป็นรายชื่อศูนย์ WWW เรียงตามลำดับขนาดของเอกสารจากขนาดใหญ่สุดไปหาเล็กสุด เอกสารอันดับแรกจะได้คะแนนเต็ม 1,000 คะแนน และเรียงลดหลั่นไปตามลำดับ ขนาดของเอกสารเป็นตัวเลขช่วยให้ผู้ใช้คาดคะเนว่าต้องใช้เวลามากน้อยเพียงใดในการนำเอกสารมาอ่าน ดังรูปที่ 5

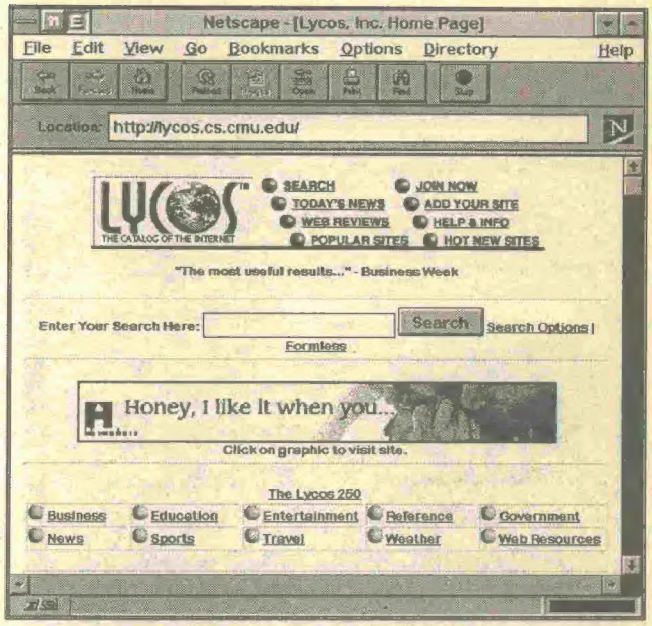
- **InfoSeek** เป็นศูนย์ค้นหาที่ให้บริการโดยคิดค่าธรรมเนียมตามการใช้งาน ซึ่งมีตั้งแต่ค่าธรรมเนียมรายครั้ง ๆ ละ 20 เซนต์ ไปจนถึงค่าบริการ 9.95 เหรียญสหรัฐต่อเดือนโดยให้ใช้งานได้ 100 ครั้ง และให้บริการแบบไม่คิดมูลค่าสำหรับการทดลองใช้ ดังรูปที่ 6

ศูนย์ InfoSeek นับเป็นศูนย์บริการค้นหา WWW ขนาดใหญ่ที่มีดรรชนีศูนย์ WWW อื่นกว่า 250,000 และยังสามารถค้นหาข้อมูลโดยป้อนเป็นคำพูดได้ เช่น "where can I play windserf"

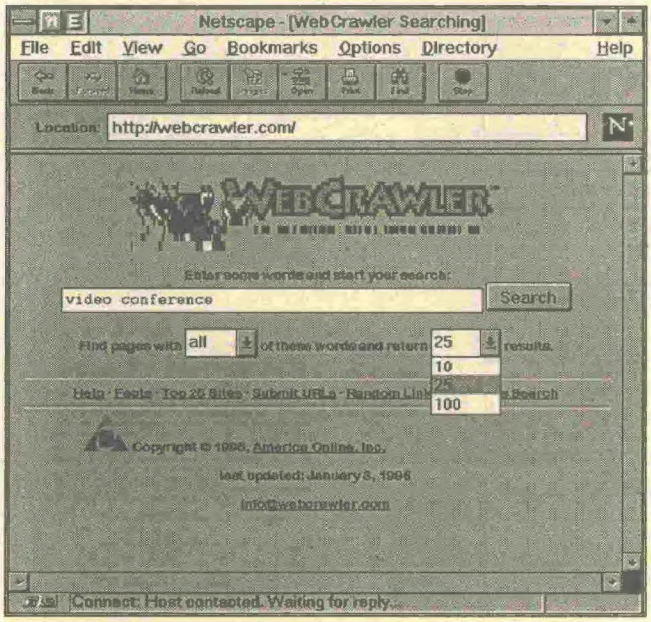
- **Lycos** เป็นแหล่งรวบรวมดรรชนีศูนย์ WWW ขนาดใหญ่อีกแห่งหนึ่งที่มีรายชื่อศูนย์กว่า 4 แสนศูนย์ ดังรูปที่ 7
- **WebCrawler** ศูนย์บริการที่มีดรรชนีศูนย์ WWW ราว 2 แสนรายการ ผลการค้นหาจะแสดงในรูปของรายชื่อศูนย์ที่มีลำดับคะแนนเรียงจากน้อยไปหามากเช่นเดียวกับ EINET ดังรูปที่ 8
- **World Wide Web Worm** (http://www.cs.colorado.edu/home/mcbryan/WWWW.html) ศูนย์บริการแห่งนี้มีผู้นิยมเข้า



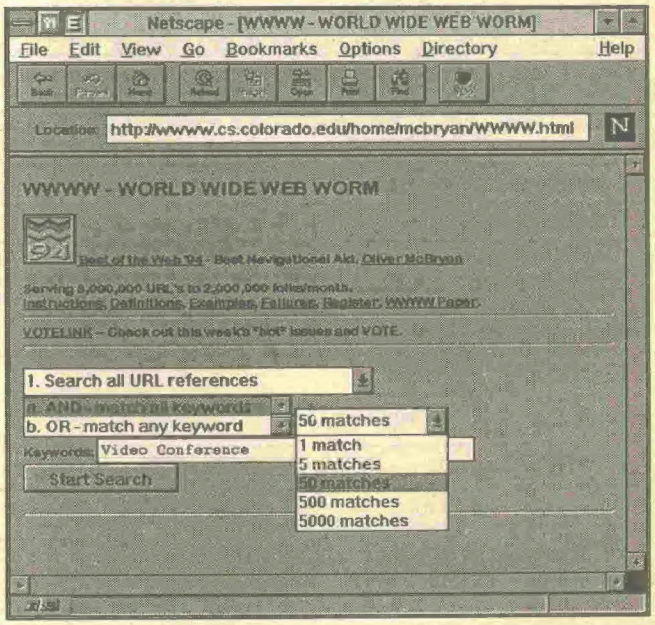
รูปที่ 6 ศูนย์ InfoSeek



รูปที่ 7 ศูนย์ Lycos



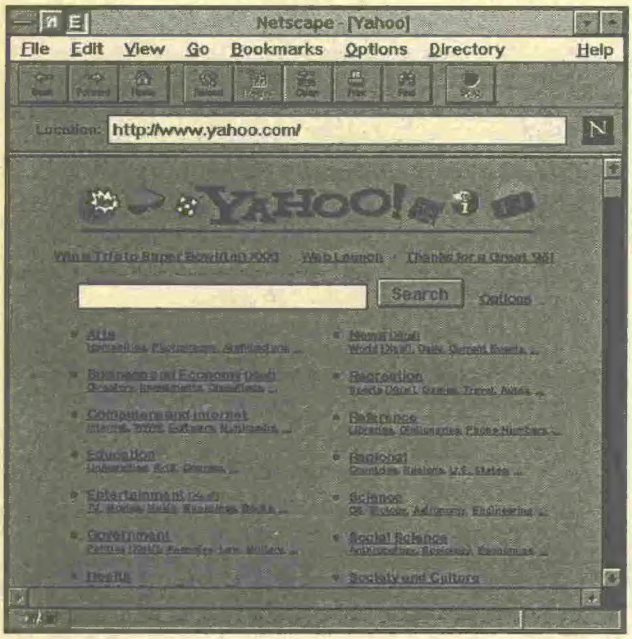
รูปที่ 8 ศูนย์ WebCrawler



รูปที่ 9 WWWW ที่โคโลราโด

ไปงานอย่างมาก และมีระบบการค้นที่สามารถกำหนดเงื่อนไขการค้นแบบ "and" และ "or" ได้ นอกจากนี้ยังสามารถกำหนดตัวเลขจำนวนศูนย์ที่ให้ค้นหาได้ด้วย ดังรูปที่ 9

● **Yahoo** เป็นหนึ่งในศูนย์บริการที่มีผู้นิยมใช้งานมาก ได้แก่ www.yahoo.com ซึ่งรวบรวมรายชื่อศูนย์ WWW ไว้กว่า 80,000 แห่ง โดยจัดแบ่งออกเป็น 14 หัวข้อ ตั้งแต่เรื่องทางศิลปะ ธุรกิจ คอมพิวเตอร์ ไปจนถึงสังคมศาสตร์ ดังรูปที่ 10



รูปที่ 10
ศูนย์ Yahoo

ไดเรคทอรี WWW

นอกจากศูนย์ WWW ที่ให้บริการค้นหาแล้ว ยังมีศูนย์ WWW อีกหลายแห่งที่จัดบริการในรูปของเมนูไดเรคทอรี กล่าวคือ เลือกหัวข้อเรื่องที่ต้องการเพื่อจำกัดขอบเขตการค้นก่อนที่จะป้อนคำสำคัญที่ให้ระบบช่วยค้นหา ตัวอย่างของศูนย์บริการเหล่านี้แสดงได้ดังตารางที่ 3

สำหรับในประเทศไทยก็มีบริการ WWW ในมหาวิทยาลัยและหน่วยงานอื่นหลายแห่ง ดังตารางที่ 4

ส่งท้าย

การสืบค้นข้อมูลในอินเทอร์เน็ตด้วย WWW เป็นที่นิยมของผู้ใช้ทั่วโลก และเป็นจุดศูนย์รวมความสนใจของคนทั่วไป นอกจากแนวคิดของไฮเปอร์มีเดียซึ่งสามารถสื่อสารข้อมูลได้ทั้งข้อความ เสียง ภาพนิ่งและภาพเคลื่อนไหวแล้ว WWW ยังช่วยเชื่อมโยงข้อมูลจากศูนย์บริการ WWW แห่งหนึ่งไปยังอีกศูนย์หนึ่งได้

พัฒนาการของ WWW ยังคงดำเนินต่อไป รวมทั้งการปรับปรุงกลไกช่วยการสืบค้นให้ทำงานได้อย่างรวดเร็วและแม่นยำยิ่งขึ้น

ตารางที่ 3 ศูนย์ค้นหา WWW ที่ให้บริการไดเรคทอรี

ศูนย์	URL
Clearinghouse for Subject-Orinted Internet Resource Guides	http://www.lib.umich.edu/chhome.html
Comprehensive List of Web Sites	http://www.netgen.com/cgi/comprehensive
Global News Network Directory	http://nearnet.gnn.com/gnn/GNNhome.html
InterNic Directory of Directories	http://www.internic.net/ds/dsdirofdirs.html
Online System Services	http://www.ossinc.net/
Yannoff Special Internet Connections List	http://www.uwm.edu/Mirror/inet.services.htmml

ตารางที่ 4 ศูนย์ WWW ในประเทศไทย

สถานที่	ศูนย์บริการ
ธนาคารแห่งประเทศไทย	http://www.npw.or.th
เนคเทค	http://www.nectec.or.th
บริษัท เคเอสซี คอมเมอร์เชียลอินเตอร์เน็ต จำกัด	http://www.ksc.net.th
บริษัท อินเตอร์เน็ตไทยแลนด์ จำกัด	http://www.inet.co.th
จุฬาลงกรณ์มหาวิทยาลัย	http://www.chulkn.chula.ac.th
มหาวิทยาลัยเกษตรศาสตร์	http://www.ku.ac.th
มหาวิทยาลัยขอนแก่น	http://www.kku.ac.th
มหาวิทยาลัยเชียงใหม่	http://www.chaingmai.ac.th
มหาวิทยาลัยธรรมศาสตร์	http://ipied.tu.ac.th
มหาวิทยาลัยมหิดล	http://www.mahidol.ac.th
มหาวิทยาลัยอัสสัมชัญ	http://www.au.ac.th
สถาบันเทคโนโลยีพระจอมเกล้าลาดกระบัง	http://www.kmitl.ac.th
สถาบันเทคโนโลยีแห่งเอเชีย	http://www.cs.ait.ac.th